图书情报工作
第63卷 第19期 2019年10月

大数据时代高校数据素养课程群构建研究*

■ 沈玖玖 徐萍 张琴 龚花萍
南昌大学管理学院信息管理系 南昌 330031

摘要：[目的/意义]大数据环境下，科学数据的管理越来越受到社会各界的重视。因此高校数据素养人才的培养日显重要，数据素养课程是对大学生进行数据素养培养的重要依据。[方法/过程]采用问卷调查和SPSS 22软件中Pearson相关性分析，对与数据素养关系较为密切的信息素养课程与数据素养之间的关系进行分析，并从中析出数据素养课程集合，再运用聚类分析对数据素养课程进行模块化分类。[结果/结论]研究结果表明信息素养相关课程可以作为数据素养的培养基础，在此基础上通过补充完善相应数据素养课程，并依据数据生命周期理论构建出包括数据意识及伦理、数据获取和数据处理、数据分析与利用课程模块的数据素养课程群。

关键词：数据素养 信息素养 课程群 相关分析 聚类分析
分类号：G252.7
DOI：10.13266/j.issn.0252-3116.2019.19.005

随着大数据时代的来临，数据形式的信息管理在科研和社会管理决策中的作用日显重要。2018年1月23日，习近平主持召开了中央全面深化改革领导小组第二次会议，会议强调要适应时代的要求，加强对数据的科学管理，积极建立数据安全机制[1]，这意味着科学数据管理已受到国家领导层的重视。因此，数据素养已成为一种亟待培养的基本能力，而高校作为培养数据素养的主要力量，应加强对学生的数据素养培养，尤其是与数据信息密切相关的专业，应适时调整数据素养教育培养目标和方案，迎合大数据时代人才需求。

纵观目前国外对数据素养教育的研究，在实施数据素养教育方面已经有较为成熟的体系。T. Koltay概述了数据密集型研究范式的主要特征，指出学术图书馆配备可以作为研究数据服务集中体现的工具[2]。E. Stephenson等通过在加州大学洛杉矶分校进行一项实验性课程，由图书管理员和数据档案管理员共同教授两次，使用教学大纲和作业，反映社会科学中的社会学研究问题和工具以及信息素养能力[3]。国内对数据素养教育的研究内容主要包括以下3个方面：①国内和国外的数据素养教育的调查、对比、启示；②高校以及高校图书馆数据素养教育的体系模型、培养机制、教育模式等方面的构建；③数据素养教育相关的综述研究。孟祥保等将国外数据素养教育实践归纳为利用Libguides建立科学数据管理资源导航、数据素养通识教育、学科数据素养教育3种模式[4]。黄如花等探讨了数据素养的概念，研究了其教学内容，论述了如何为不同机构提供数据素养教育且对受众进行分类[5]。此外，魏来等[6]从教学对象、教学途径、教学内容等方面出发，研究我国各大高校的信息素养培养建设情况，对我国课程内容建设提出建议。

通过阅读大量文献发现，目前关于数据素养教育的研究仍停留在宏观模式研究阶段，尚未涉及数据素养相关课程设置方面，但数据素养教育需要具体的实施依据。因此，本文通过咨询专家意见，析出与数据素养密切相关的信息管理与信息系统专业（以下简称“信管专业”）的信息素养相关课程，通过问卷调查和Pearson相关性分析方法，提炼出其中对数据素养影响显著的相关课程，同时根据大数据时代的要求，融入相关最新的数据管理方面的新知识和新内容，并根据数据生命周期理论，形成独立的数据素养课程体系，为高校大学生数据素养通识教育提供实施依据。其中，对数据素养课程的模块化划分将有助于各培养机构有的放矢

* 本文系2017年江西省高等学校教学改革研究项目“大数据时代大学生数据素养教育课程群构建研究——以信息管理与信息系统专业为例”（项目编号：JXJG-17-1-55）和“大数据视阈下大学生数据素养成长规律与培养模式研究”（项目编号：JGLX-18-1-55）研究成果之一。

作者简介：沈玖玖（ORCID：0000-0002-9415-4363），教授，硕士，硕士生导师，E-mail：1435657047@qq.com；徐萍（ORCID：0000-0001-8800-6455），本科生；张琴（ORCID：0000-0002-4862-8775），本科生；龚花萍（ORCID：0000-0003-4601-7628），教授，硕士，硕士生导师。

收稿日期：2018-12-23 修回日期：2019-03-04 本文起止页码：66-74 本文责任编辑：刘远颖

地根据各专业的具体要求实施个性化培养计划。

1 理论基础

1.1 数据素养的内涵

现有相关研究中，对于数据素养的概念未给出标准定论，数据素养多被研究者认为是信息素养的延伸。目前学术界引用率较高的定义分为两种：一是 E. Stephenson 等[3]将数据素养定义为"能够找到、评价和有效合理使用数据的能力"；二是 K. Hogenboom 等认为数据素养是阅读、解释、分析、批判性思考数据以及将数据作为证据的能力[7]。此外，其他学者也对数据素养进行了定义，例如，J. R. Carlson 等认为数据素养基础能力包括数据的格式、数据管理、数据收集、数据处理应用、数据可视化和数据伦理在内的 12 项核心能力[8]。胡卉等认为数据素养是指在合理规范下获取数据、客观地选择和评估数据，能够管理、处理、利用数据等[9]。隆茜认为高校师生数据素养包括数据意识、数据获取、数据处理与分析、数据交流、数据评估及数据道德[10]。

根据前人的研究，数据素养课程的建设将围绕数据意识、数据伦理、数据获取、数据处理、数据分析与利用 5 个维度进行。其中，数据意识是指了解数据重要性和数据获取途径，是一种主观感受；数据伦理是指与数据相关的规则与法律政策，例如标注数据来源、保证数据真实性等；数据获取能力是指精通获取一手数据和二手数据的方法，例如进行问卷调查以及使用多种数据检索工具；数据处理能力是指运用各种数据处理工具和软件（Ucinet、SPSS、AMOS 等）对采集的数据进行可视化处理并以图表的形式展示出来的能力；数据分析能力是指熟练使用多种数据分析软件，具备选择有效的数据分析方法并挖掘其内在价值的能力。

1.2 课程群内涵

课程群是指课程体系中具有某种性质或担负类似功能的几门课程的小集群[11]，是基于知识体系构筑的有机课程体系模块，高校课程群有"文理学科通识课程群、学科通识课程群、专业基础课程群、各专业课程群"四级体系[12]。笔者认为，数据素养课程群是指以培养大学生的数据素养为目标，由影响数据素养的相关课程聚集而成的有机集合。它既可以作为强化大学生数据素养教育的依据，也可以为有数据素养要求的相关专业调整培养计划提供参考。

1.3 数据素养与信息素养的关系

孟祥保等[4]提出数据素养与信息素养类似，认为数据素养教育可以视为信息素养教育在大数据时代的拓展和细化[13]。张进良等[14]认为数据素养与信息素养有着不容忽视的密切关系，数据素养是信息素养的必要组成部分。数据素养是信息素养在大数据背景下的创新性发展，是信息素养的核心组成部分[15]。J. C. Prado 等[16]将数据素养定义为信息素养的组成部分。张静波[17]认为数据素养是信息素养的子集。数据素养教育是传统信息素养教育的进一步提高和深化。据此，笔者也认为，信息包含各种类型的信息，数据型信息是信息中的一种类型，信息素养应包括数据素养。因此信息素养课程一定程度上可能会对数据素养的培养造成影响，因此数据素养课程群可以参考信息素养课程进行构建。

1.4 信息素养相关课程设置及释义

我国的信息管理与信息系统专业是在教育部 1998 年 7 月颁布的《普通高等学校本科专业目录和专业简介》[18]之后开设的，目前在北京大学、武汉大学、南京大学、中山大学、华东师范大学、黑龙江大学、华中师范大学、南昌大学等高校均有设立。其中，以较有代表性的武汉大学为例，其信息管理与信息系统专业的培养目标表明，该专业培养具有信息管理学基础、具备较高的信息素养、掌握系统思想以及信息系统设计与管理方法等方面的知识与能力，能在国家各部门从事信息管理以及信息系统分析、设计、实施、管理和评价等方面工作的高级复合、创新型人才。信息素养是该专业学生的核心能力，相关高校的信息管理与信息系统专业培养计划的专业课程中设置了信息素养的相关课程，笔者在咨询了该专业的 6 位教师和专家的前提下，从部分高校的信息管理与信息系统专业课程中提炼出与信息素养密切相关的课程，见表 1。据笔者分析，这些不同高校的信息素养课程设置虽有些许出入，但总体近似度较高。为了保证收集数据的便利性和可靠性，课题组选择所在的大学管理学院信息管理与信息系统专业作为调研对象。该专业培养方案是按照教育指导委员会培养大纲的精神制定的，先后经过多次修订，目前使用的培养方案属 2016 年版。其中培养方案中的专业课程主要包括经济管理类、计算机类、信息类三大模块，去除经济管理类及部分纯计算机语言类等课程之后的主要课程及详情见表 2，笔者将通过实证调研及相关性分析，探究这些相关课程在数据素养培养方面的影响效果，检验出其中与数据素养密切相关的课程，并以此为基础构建适合数据素养教育的课程集合。

表 1 高校信息管理与信息系统信息素养相关课程

大学	信息管理与信息系统专业的部分专业课程
北京大学	信息资源管理与知识管理、信息存储、信息组织、信息检索、传播与数据挖掘、文献目录学、政府与企业信息化管理、信息资源产业、网络技术与应用、情报分析与咨询、信息政策与法规
武汉大学	信息管理概论、高级语言程序设计、数据结构、数据库系统原理、计算机网络、信息资源管理、智能信息系统、信息组织、信息分析与预测、信息检索、管理信息系统、信息计量学、知识管理等
南京大学	信息资源管理导论、信息技术基础、信息组织(上、下)、信息检索(中文、西文)、计算机网络、信息分析、数据库原理与应用、数据结构、管理信息系统、信息传播技术、计算机系统与系统软件、程序设计语言等
中山大学	信息管理概论、信息组织、信息检索、信息分析与决策、管理信息系统、竞争情报研究、数据结构、程序设计、数据库系统、计算机网络等
华东师范大学	数据库、信息管理概论、信息检索、信息处理、信息分析、信息咨询与服务、管理信息系统
黑龙江大学	应用数理统计、数据结构、数据库计算机网络、信息管理学、信息组织、信息存储与检索、信息系统分析与设计、计算机信息检索、信息经济学
华中师范大学	管理信息系统、数据库原理、计算机网络、数据结构、信息管理学、信息采集学、信息组织、信息检索、信息服务与信息用户研究、数据库应用、信息政策与法规、信息分析与预测
南昌大学	现代信息技术基础、数据结构、信息资源管理、数据库原理、统计学、信息存储与检索、信息计量学、信息分析与预测、数据挖掘、信息可视化、网络信息安全、信息系统分析与设计、知识产权保护、情报工程学

注:数据来源于各大高校官网

2 实证研究

2.1 问卷调查

本次调查主要针对南昌大学管理学院信息管理与信息系统专业的 2015 级至 2018 级的在校生,采用线下发放纸质问卷的方式进行。调研时间安排在 2018 年 10 月 17 日至 11 月 15 日,问卷样本总共发放 210 份,回收 200 份,有效问卷 192 份。

表 2 南昌大学信息管理与信息系统专业信息素养相关课程及释义

课程	课程性质	授课模式	学期安排	课程性质及简介
现代信息技术基础	基础课	理论 + 实验	一	主要介绍计算机基本知识、网络技术基础及应用。旨在培养学生信息意识
数据结构	基础课	理论 + 实验	三	讨论各种数据结构的逻辑结构、存储结构及有关操作的算法。旨在培养学生信息获取及处理能力
信息资源管理	专业课	理论 + 实验	三	介绍信息资源的类型、采集、组织、检索及开发利用与信息管理。旨在培养学生信息意识、信息道德和信息获取能力
数据库原理	基础课	理论 + 实验	三	介绍数据查询、数据库的完整性控制、安全性机制和数据库设计的方法等。旨在培养学生综合信息素养
统计学	基础课	理论	四	介绍数据的收集、整理、分析及综合评价方法等。旨在培养学生信息获取及处理能力
信息存储与检索	专业课	理论 + 实验	五	介绍重要的中、外文检索系统的检索方法及网络信息检索方法。旨在培养学生信息获取能力
信息计量学	专业课	理论 + 实验	五	介绍信息计量学主要的定量分析方法。旨在提高学生信息处理及分析能力
信息分析与预测	专业课	理论	五	介绍信息分析与预测的基本理论和方法、信息分析与预测的活动流程。旨在培养学生信息分析能力
数据挖掘	专业课	理论 + 实验	五	介绍数据挖掘的基本概念,原理、方法和技术。旨在培养学生信息处理能力、信息分析及利用能力
信息可视化	专业课	理论	五	介绍图形图像技术并对各类信息和数据进行可视化。旨在培养学生信息处理能力
网络信息安全	专业课	理论 + 实验	六	综合介绍信息安全和网络安全两个方面。旨在培养学生信息伦理意识
信息系统分析与设计	专业课	理论 + 实验	六	介绍信息系统中信息系统建设的规划、开发、维护和管理过程。旨在培养学生综合信息素养
知识产权保护	专业课	理论	六	介绍我国关于知识产权方面的相关法律、相关政策及国际上关于这方面的相关法律规定。旨在使学生了解信息伦理

2.1.1 问卷设计 问卷分为 3 个部分:①用户基本信息调研,设置性别、年级题项;②数据素养能力调研,设置反映数据素养能力题项;③信息素养相关课程调研,设置信息素养具体课程的题项(包括课程学习掌握程度和实验课的辅助作用)。后两部分调研采用矩阵量表的形式,量表采用里克特 5 点式:第二部分等级划分为 1 = 非常不符合,2 = 比较不符合,3 = 一般符合,4 = 比较符合,5 = 非常符合;第三部分等级划分为 1 = 非常差,2 = 比较差,3 = 一般,4 = 比较好,5 = 非常好。具体内容如表 3 所示:

沈玖玖，徐萍，张琴，等.大数据时代高校数据素养课程群构建研究[J].图书情报工作,2019,63(19):66－74.

表3　问卷调研题项设置

项目	指标	编号	题项
数据素养能力	数据意识	A	知道识别和区分数据的格式、类型和特征的方法
			知道多渠道查找所需数据
			知道采取合适的管理方式与工具使得数据便于使用
			经常通过数据加密、数据访问权限控制来保证数据私密性
	数据道德	B	尊重他人的数据并根据科学或行业规则使用数据
			自觉遵守相关法规,合法获得数据和使用信息产品
			遵守数据收集、使用中所涉及的道德和伦理
	数据获取	C	善于全面准确地识别并收集所需要的所有类型的数据
			善于利用合适的检索工具查找所需要的数据
			善于选择有效的收集方法(调查、访谈、实验等)来收集数据
	数据处理	D	善于使用一些常见的数据处理工具(如SPSS、Excel等)
			精通简单的可视化工具(如MySQL、Excel、Python等)
			善于规避数据呈现时模棱两可的表述情形
	数据分析与利用	E	善于选择正确有效的数据分析方法(如线性回归分析、方差分析)
			熟练使用一种或多种数据分析软件(如SPSS、MATLAP等)
			善于通过构建模型等方式深入挖掘数据隐藏信息
			善于依据数据分析的结果做出正确的结论和决策
			善于正确有效地利用数据来分析并解释现实问题
信息素养相关课程	现代信息技术基础	F	本课程的学习和掌握程度以及实验课的辅助作用
	数据结构		本课程的学习和掌握程度以及实验课的辅助作用
	信息资源管理		本课程的学习和掌握程度以及实验课的辅助作用
	数据库原理		本课程的学习和掌握程度以及实验课的辅助作用
	统计学		本课程的学习和掌握程度以及实验课的辅助作用
	信息存储与检索		本课程的学习和掌握程度以及实验课的辅助作用
	信息计量学		本课程的学习和掌握程度以及实验课的辅助作用
	信息分析与预测		本课程的学习和掌握程度
	数据挖掘		本课程的学习和掌握程度以及实验课的辅助作用
	信息可视化		本课程的学习和掌握程度
	网络信息安全		本课程的学习和掌握程度以及实验课的辅助作用
	信息系统分析与设计		本课程的学习和掌握程度以及实验课的辅助作用
	知识产权保护		本课程的学习和掌握程度

2.1.2　样本概况　对192份有效问卷中个人信息进行统计发现,女性为87人,所占比例为45.31%;男性为105人,所占比例为54.69%,性别和年级分布比例基本均匀,样本显示有较强的代表性,具体概况如表4所示:

表4　样本概况

名称	选项	频数	百分比(%)
性别	女	87	45.31
	男	105	54.69
年级	2015级	40	20.83
	2016级	59	30.73
	2017级	44	22.92
	2018级	49	25.52

2.1.3　信效度分析

(1)信度。本次数据共涉及5个维度,数据意识、数据伦理、数据获取、数据处理、数据分析和利用5个维度分别用A-E表示。采用SPSS 22分析软件,用克隆巴赫信度系数(Cronbachα系数)去测量数据的信度质量水平。α系数值高于0.8,则说明信度高;α系数介于0.7－0.8之间,则说明信度较好;α系数介于0.6－0.7,则说明信度可接受;α系数小于0.6,说明信度不佳。本次调查问卷信度综合整理见表5,6个维度的α系数值均高于0.6,最小为0.664,说明本次数据的信度质量水平较好,研究数据真实可靠。

(2)效度。本文采用KMO检验和Bartlett's球体检验进行样本数据的效度分析,当KMO检验系数取值

图书情报工作
第 63 卷 第 19 期 2019 年 10 月

表 5 信效度结果

维度	测量指标	Alpha 系数	项删除的 Alpha 系数	KMO 测度
整体	所有指标	0.886	—	0.808
数据意识	A1	0.664	0.549	0.693
	A2		0.562	
	A3		0.465	
	A4		0.76	
数据伦理	B1	0.783	0.56	0.513
	B2		0.344	
	B3		0.487	
数据获取	C1	0.822	0.713	0.677
	C2		0.663	
	C3		0.755	
数据处理	D1	0.833	0.794	0.632
	D2		0.639	
	D3		0.652	
数据分析与利用	E1	0.795	0.759	0.776
	E2		0.652	
	E3		0.616	
	E4		0.549	

大于 0.5，Bartlett's 球体检验的卡方统计值显著性水平 P 值小于 0.05 时，各潜变量均具有收敛度和区别度，表示该问卷具有良好的建构效度。首次对 6 个维度、17 个量表题进行分析，结果见表 5。KMO 检验系数整体取值为 0.808，Bartlett's 球体检验的卡方统计值显著性水平 P 值严格小于 0.05，在理想范围内，所以本研究假设模型的数据样本具有效度。

2.2 结果分析

2.2.1 各年级数据素养水平差异分析　为了测定各年级数据素养的水平，笔者首先对各年级的数据素养各维度的平均值进行比较，其折线图见图 1。分析各年级数据素养各维度总体趋势，发现随着年级上升，大学生数据素养水平呈现上升趋势，尤其是对数据意识、数据获取两个维度影响较大。说明随年级升高，学生学习的相关课程在增多，数据素养相关知识也在逐渐丰厚，但各维度水平提升程度有限，说明所设置的课程在培养学生数据素养方面有一定的效果，但影响程度有限，提示我们构建数据素养课程群时需要加强与数据伦理、数据处理和数据利用能力相关的课程设置。

由图 1 所展示的各年级数据素养水平差异可以发现，随着年级的上升，高年级相较低年级虽然所学相关课程更多，但与低年级相比，数据伦理维度等级没有明显差异，各年级的数据伦理平均分均维持在 3.2 – 3.4 之间，总体均处于中上等水平，说明数据伦理并没有通过相关课程的学习获得提升。笔者认为这种结果显示

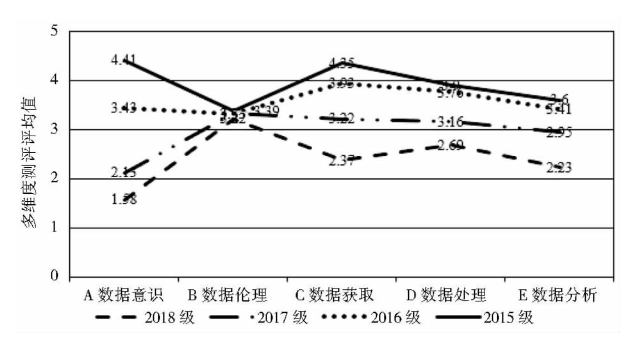


图 1 2015 级 – 2018 级数据素养多维度测评平均值

数据伦理的形成可能来自于课程之外的学习或训练。因此相关课程在培养学生的数据伦理方面存在较大的完善空间。此外，数据意识、数据获取能力随年级的上升明显增强，数值达到等级 4，意味着该专业的相关课程对学生的数据意识的培养及数据获取能力的提高有较好的培养效果。而数据处理和数据分析能力，随年级上升，没有明显提升，数值均未达到等级 4，则提示在构建数据素养课程时应考虑完善与数据处理和数据分析能力相关的课程。

2.2.2 实验课程配套　笔者根据问卷调研，得到学生对实验课在辅助对应课程知识掌握中的程度的评分，其评分的平均值见图 2。由图 2 可知，配套实验课对学生掌握对应课程知识会产生助推作用，且程度都在中等偏上的水平，提示数据素养相关课程也有必要设置对应的实验课，但数据显示实验课的助推效果均未达到等级 4，说明该专业的实验课设置有待改进，例如可增加实验设施投入，改善实验环境，增加实验课时等。

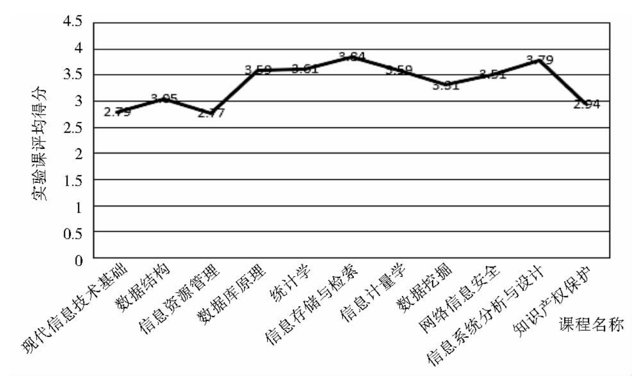


图 2 实验课设置效果评分

2.2.3 数据素养与相关课程间的相关性分析　基于 Pearson 系数广泛用于度量两个变量之间的相关程度，文章通过 SPSS 22 软件进行 Pearson 相关性分析，研究信息素养相关课程与数据素养维度之间的相关关系。Pearson 相关系数值可以反映相关性强度，具体情况见表 6[19]。相关分析结果见表 7。从中可以发现，大多数

课程均与相关数据素养维度间存在较高的相关性，仅有“现代信息技术基础”和“信息资源管理”课程与数据素养维度之间的最大相关系数分别只达 0.179 和 0.192，不足 0.2。为了使得数据素养课程更加精炼，笔者把这种相关性较弱的课程排除，留下剩余的 9 门课程，作为数据素养课程。

表 6　相关系数强度

相关系数	相关性强度
|R| ≤0.2	无相关或弱相关
0.2 ＜ |R| ≤ 0.6	相关
0.6 ＜ |R| ≤ 0.8	中度相关
0.8 ＜ |R| ≤ 1	强相关

表 7　数据素养各维度与课程相关性

课程名称	数据意识	数据伦理	数据获取	数据处理	数据分析与利用
现代信息技术基础	0.023	－0.245	0.192	0.136	0.126
数据结构	0.025	－0.04	0.535	0.379	0.164
信息资源管理	0.179	0.114	0.100	0.001	0.144
数据库原理	0.023	0.092	0.442	0.401	0.192
统计学	0.149	－0.249	0.579	0.41	0.328
信息存储与检索	0.453	0.337	0.138	0.125	－0.005
信息计量学	－0.09	0.022	0.419	0.46	0.343
信息分析与预测	0.205	0.272	0.263	0.499	0.568
数据挖掘	0.123	－0.197	0.485	0.387	0.548
信息可视化	0.122	－0.162	0.282	0.478	0.446
网络信息安全	0.497	0.203	0.435	－0.023	－0.042
信息系统分析与设计	0.353	0.351	0.279	0.419	0.351
知识产权保护	0.658	0.22	0.012	0.056	0.058

2.2.4　数据素养相关课程的聚类分析　为了进一步明确以上所确定的数据素养课程与数据素养各维度间的密切程度，笔者将上文测算的相关系数作为自变量，使用 SPSS 22 软件对以上 11 种课程样本数据进行聚类分析。将对数据素养各维度影响相似的课程聚集在一起，进行模块化划分。而聚类分析只是将距离相对最近的课程聚集在一起，并不能说明一门课程不会同时影响两个维度，例如，“信息存储与检索”课程在相关系数分析中，对数据的获取也会具有一定的影响，但是其与数据意识更为相关，则把该门课程归为数据意识类课程。聚类方法采用最小距离法（Between-group-sinkage），距离测试选择欧式距离法（Euclidean Distance），聚类结果见图 3。

根据聚类结果，对各课程基于数据素养维度进行模块划分。笔者认为数据意识与数据伦理是数据素养的基础维度，也即是数据获取、数据处理及分析利用的前提，且两类素质相关的课程有很大相通性，笔者将这两个维度的相关课程统一归为一类，即数据素养意识及伦理培养模块。且由图 3 可知，数据获取和数据处理相关课程的距离较近，表明数据获取与数据处理具有相通性，所以笔者将两个维度的相关课程归为一类，命名为数据操作模块。总之，笔者将聚类结果分为 3 个模块：数据意识及伦理模块、数据操作模块和数据分析与利用模块。根据各类所包含的课程，总结各模块特点如下：第一类，数据意识及伦理模块。该课程模块

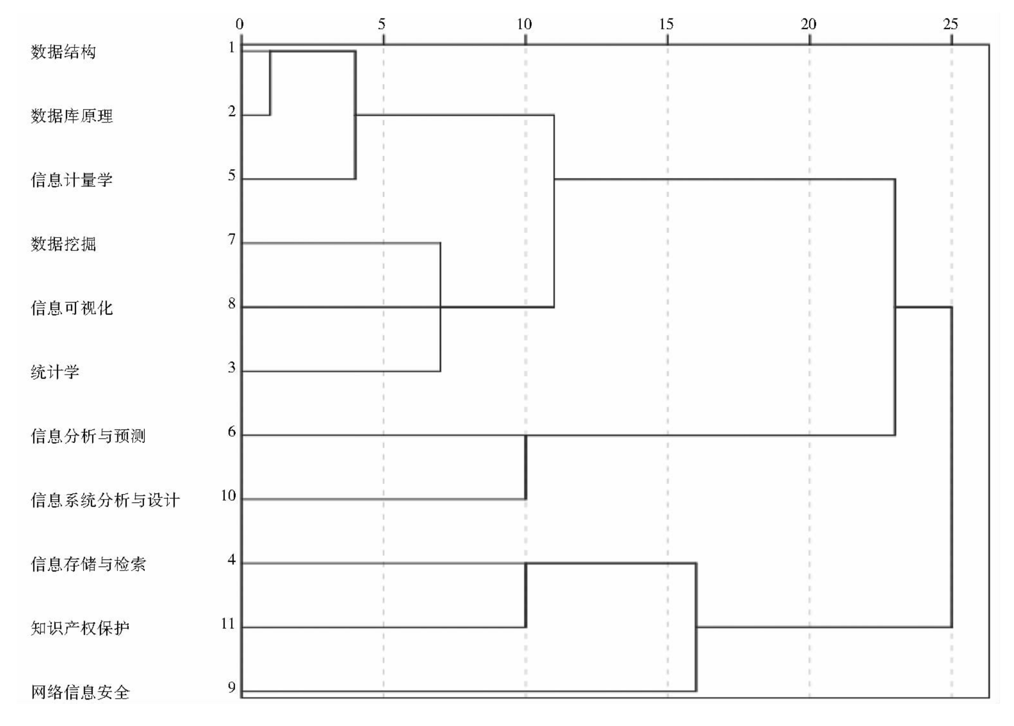


图 3　数据素养相关课程聚类结果

旨在对大学生进行数据意识、数据管理、数据道德、数据安全和大数据基础知识等方面的培养，主要课程包括“信息存储与检索”“知识产权保护”“网络信息安全”等。第二类，数据操作模块。该课程模块旨在培养学生数据收集存储能力和数据加工以及数据呈现方法的技能，其中包括相关数据收集、数据处理软件的学习，主要课程有“数据结构”“数据库原理”“信息计量学”“数据挖掘”“信息可视化”“统计学”等。第三类，数据分析与利用模块。该课程模块旨在培养学生数据分析和数据价值发现能力，了解各种方法使用的场景并根据分析结果得出结论，包括“信息分析与预测”“信息系统分析与设计”等课程。

3 数据素养课程群构建

3.1 数据素养课程的补充

鉴于图 1 所示的该专业各年级学生数据伦理维度等级没有明显差异，该专业学生数据处理和数据分析能力提升幅度也不太明显，相关课程的设置还存在较大的提升和完善空间，因此，笔者基于前人研究并参考其他高校信息管理专业的培养方案（见表 1）中信息素养相关课程的开设情况，在咨询了本校相关专业的 6 位教师和专家之后，针对前文所构建的课程模块进行相应的补充和完善，增加数据素养相关课程设置，如表 8 所示：

表 8 数据素养补充课程表

数据素养维度	课程	课程简介	开设高校	数据来源
数据伦理	数据素养:数据使用者指南	主要介绍数据来源标注、数据识别	耶鲁大学	文献[5]
	信息政策与法规	主要介绍信息方面相关的法律法规	北京大学、华中师范大学	各大高校官网（见表 1）
数据处理	用 Python 玩转数据	主要介绍基于机器语言的数据获取与表示、数据可视化相关知识	南京大学	
	信息组织	主要介绍各种数据有序化的处理方式	北京大学	
数据分析与利用	R 语言与数据分析	主要介绍数据分析方法（参数估计、假设检验、多元统计分析方法及贝叶斯统计分析）以及提供大量案例分析	南京大学	
	社会调查和统计分析	主要介绍社会调查方案、数据分析应用方法	西南交通大学	

此外，基于前文实验配套设置分析，实验课设置对掌握课程知识有较大的辅助作用，笔者认为在不影响基础课时的前提下，可结合理论课程的特点，适当开设相应配套实验课，提高学生的数据操作能力。此外，还应借助实验室操作技能培训方式对学生的数据管理实践能力进行训练。同时，学校应建立设备先进的数据素养操作培训实验中心，为数据素养教育相关课程的理论教学提供配套实验资源支持，使学生的数据素养在实践教学中不断得到改进和稳步的提升。

3.2 数据素养课程群结构模式

数据生命周期是指数据从收集（或形成）、加工、保存、传播、检索、存取与利用，到消失或不再被利用的过程[17]。在数据生命周期中，从数据获取到数据利用存在一定的逻辑关系，即数据处理受到数据收集的驱动，数据处理服务于数据分析，数据分析的目的是挖掘数据价值。而数据意识与伦理是数据获取—处理—分析过程的基本前提，即获取数据首先需要建立在数据意识的基础上，同时必须遵守数据伦理，如必须通过正规渠道，数据处理时不能随意篡改数据，数据利用时需注明数据来源、引用等，否则后续的数据处理、分析与利用都不可能正常进行或受到影响。因此数据素养课程群的构建应遵循数据生命周期的逻辑关系进行课程的有序安排。数据素养课程的设置、内容、授课模式及排序等见表 9。

4 结语

在大数据时代，人们对数据的需求更为迫切。高校作为教育的主阵地，理应加强数据素养教育，培养学生认识和挖掘数据价值的能力。本文以信息管理专业信息素养相关课程为基础，通过相关性分析方法，重点研究信息素养相关课程与数据素养的相关性，提炼出与数据素养相关的课程。此外，通过聚类分析将数据素养相关课程进行模块化划分，构建数据素养课程群，并基于目前课程设置的不足，对数据素养课程群进行完善。最终，根据数据的生命周期理论构建出数据素养课程群。因此，笔者就高校数据素养课程群建设与实施有以下建议：①各专业在采用本课程群实施数据素养教育时，需考虑学生的专业学科背景和基础条件，如文理工科专业差异性较大，各专业学科可以结合专业需要在本课程群基础上进行适当调整，构建适合自身专业学生的数据素养课程群；②数据素养课程结构还应当随着时代需求的变化而实时更新。

表 9　数据素养课程群

类别	课程	授课模式	课程主要教学内容	数据生命周期	
				涉及数据素养维度	顺序安排
数据意识及伦理模块	数据素养:数据使用者指南	理论＋实验	数据来源标注、数据识别	数据意识	
	知识产权保护	理论	数据道德、产权保护	数据伦理	
	网络信息安全	理论＋实验	信息安全、网络安全		
	信息政策与法规	理论	数据管理、数据组织、数据存储		
	信息存储与检索	理论＋实验	数据收集、组织、存储		
数据操作模块	数据结构	理论＋实验	数据结构的逻辑结构,存储结构及有关操作的算法	数据获取	
	数据库原理	理论＋实验	数据查询、安全性机制和数据库设计的方法		
	信息计量学	理论＋实验	文献相关定量分析		
	统计学	理论＋实验	数据的收集、整理、分析及综合评价方法等	数据加工	
	数据挖掘	理论＋实验	数据挖掘的基本概念,原理、方法和技术	数据保存	
	信息可视化	理论＋实验	图形图像技术并对各类信息和数据进行可视化		
	用 Python 玩转数据	理论＋实验	数据获取与表示、Python 相关学习,数据可视化		
	信息组织	理论＋实验	各种数据有序化的方法		
数据保存数据分析与利用模块	R 语言与数据分析	理论＋实验	数据分析方法(参数估计、假设检验、多元统计分析方法及贝叶斯统计分析)和案例分析	数据分析与再利用	
	信息分析与预测	理论＋实验	信息分析方法运用		
	信息系统分析与设计	理论＋实验	信息系统建设的规划、开发、维护和管理过程		
	社会调查和统计分析	理论＋实验	社会调查方案、数据分析应用		

本研究还存在不足之处,即在对数据素养课程群进行完善的过程中,未对新添加的课程与数据素养各维度能力进行相关性分析,无法确定其对数据素养的影响效果,待本研究成果实际实施之后,在后续研究中,课题组将开展这些新补充的课程与数据素养的相关性分析,从而不断完善数据素养课程群的建设研究。

参考文献:

[1] 习近平:思想再解放 改革再深入 工作再抓实推动全面深化改革在新起点上实现新突破[EB/OL]. [2018－08－09]. http://www.xinhuanet.com/politics/2018-01/23/c_1122303539.htm.

[2] KOLTAY T. Facing the challenge of data-Intensive research: research data services and data literacy in academic libraries[M]. Eger: Innovation in libraries and information services, 2016.

[3] STEPHENSON E, CARAVELLO P S. Incorporating data literacy into undergraduate information literacy programs in the social sciences: a pilot project[J]. Reference services review, 2007, 5(4): 525－540.

[4] 孟祥保, 李爱国. 国外高校图书馆科学数据素养教育研究[J]. 大学图书馆学报, 2014, 32(3): 11－16.

[5] 黄如花, 李白杨. 数据素养教育:大数据时代信息素养教育的拓展[J]. 图书情报知识, 2016 (1): 21－29.

[6] 魏来, 王思明. 我国高校图书馆数据素养课程内容构建研究[J]. 情报资料工作, 2018 (6): 90－95.

[7] HOGENBOOM K, HAYSLETT M. Pioneers in 24 the wild west: managing data collection[J]. Library and the academy, 2017,17(2):295－311.

[8] CARLSON J R, FOSMIRE M, MILLER C, et al. Determining data information literacy needs: a study of students and research faculty[J]. Libraries and the academy, 2011 (2) :629－657.

[9] 胡卉, 吴鸣. 浅析信息素养教育为数据素养教育带来的启示[J]. 图书馆理论与实践, 2016 (11): 9－14.

[10] 隆茜. 数据素养能力指标体系构建及高校师生数据素养能力现状调查与分析[J]. 图书馆, 2015 (12): 51－56.

[11] 朱沙. 电子商务专业信息技术课程群规划设置研究[J]. 电脑知识与技术, 2018, 14(24): 271－272.

[12] 阳志标, 张丹. 基于课程群的高校信息素养教育课程体系构建研究[J]. 中国管理信息化, 2017, 20(24): 247－248.

[13] 孟祥保, 常娥, 叶兰. 数据素养研究:源起、现状与展望[J]. 中国图书馆学报, 2016, 42(2): 109－126.

[14] 张进良, 李保臻. 大数据背景下教师数据素养的内涵、价值与发展路径[J]. 电化教育研究, 2015, 36(7): 14－19.

[15] MANDINACH E B, CUMMER E S. A systemic view of implementing data literacy in educator preparation[J]. Educational research, 2013, 42(1): 32－37.

[16] PRADO J C, MARZAL M A. Incorporating data literacy into information literacy programs: core competencies and contents[J]. Libri, 2013, 63(2):123－134.

[17] 张静波. 大数据时代的数据素养教育[J]. 科学, 2013, 65(4): 29－32.

[18] 王晓东, 吴雅琴, 王呼生. 以信息化人才培养为核心的信息管理与信息系统专业建设探究[J]. 内蒙古医学院学报, 2012(8): 539－544.

[19] 付志刚. 统计学[M]. 上海:经济管理出版,2015.

作者贡献说明:

沈玖玖:负责论文选题规划,修改论文;
徐萍:撰写与修改论文;
张琴:问卷收集与统计;
龚花萍:修改论文。

Research on the Construction of University Data Literacy Curriculum Group in the Era of Big Data

Shen Jiujiu　Xu Ping　Zhang Qin　Gong Huaping
Department of Information Management, School of Management, Nanchang University, Nanchang 330031

Abstract: [**Purpose/significance**] In the era of big data, the management of scientific data has received more and more attention from all walks of life. Therefore, the cultivation of data literacy talents in colleges and universities is becoming more and more important. The data literacy curriculum is an important basis for the data literacy training of college students. [**Method/process**] Using the questionnaire survey and Pearson correlation analysis method in SPSS. 22 software, the article analyzed the relationship between information literacy curriculum which is closely related to data literacy and data literacy, and then refined the data literacy curriculum collection. Finally, clustering analysis was used to classify the data literacy curriculum. [**Result/conclusion**] The research results show that courses related to information literacy can be used as the basis for data literacy. On this basis, by supplementing and perfecting the corresponding data literacy courses, the article constructs the data literacy course group, including data awareness and ethics course module, data acquisition and data processing course module and data analysis and utilization course module.

Keywords: data literacy　information literacy　course group　related analysis　cluster analysis